

Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-05 16:38:07

PAGE 1

REFERENCE NO: 292

This contribution was submitted to the National Science Foundation as part of the NSF CI 2030 planning activity through an NSF Request for Information, https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf17031. Consideration of this contribution in NSF's planning process and any NSF-provided public accessibility of this document does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government. The content of this submission is protected by the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>).

Author Names & Affiliations

- Shaowen Wang - University Consortium for Geographic Information Science (UCGIS)

Contact Email Address (for NSF use only)

(Hidden)

Research Domain, discipline, and sub-discipline

Computational and Data Sciences, CyberGIS, Environment and Geography, Geospatial Sciences and Technologies, High-Performance Computing

Title of Submission

Synergistically Advancing Cyberinfrastructure and CyberGIS to Revolutionize Geospatial Discovery and Innovation

Abstract (maximum ~200 words).

CyberGIS represents an interdisciplinary field combining advanced computing and cyberinfrastructure, geographic information science and systems (GIS), spatial analysis and modeling, and a number of geospatial domains (e.g., emergency management, smart cities, and the nexus of food, energy, and water systems) to enable broad scientific and technological advances. It has also emerged as new-generation GIS based on holistic integration of high-performance and distributed computing, data-driven knowledge discovery, visualization and visual analytics, and collaborative problem-solving and decision-making capabilities. The growing importance of cyberGIS is reflected by increasing calls for solutions to bridge the significant digital divide between advanced cyberinfrastructure and geospatial communities in the big data era. This report discusses challenges and opportunities for synergistically advancing cyberinfrastructure and cyberGIS to revolutionize geospatial discovery and innovation through interdisciplinary approaches.

Question 1 Research Challenge(s) (maximum ~1200 words): Describe current or emerging science or engineering research challenge(s), providing context in terms of recent research activities and standing questions in the field.

Geospatially heterogeneous and multi-scale changes across the globe, such as population growth, competing land uses, interdependencies of critical infrastructures, and accelerated resource and environmental degradation, have posed grand challenges ranging from emergency management and critical infrastructure planning to health and wellness. To tackle these challenges, which exhibit significant geo and spatial complexity, requires the consideration of interactions of spatial patterns and their driving processes across a number of spatial and temporal scales through geospatial analysis, modeling, and synthesis. Such geospatial approaches involve combining geospatial big data, analytics and models to form novel problem-solving approaches enabled by cyberGIS.

Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-05 16:38:07

PAGE 2

REFERENCE NO: 292

The development and use of geographic information science and systems (GIS) – computerized approaches to capturing, transforming, managing, analyzing and visualizing geospatial data – have grown immensely for the past several decades. It is clear that this growing trend will persist into the foreseeable future driven by massive needs in numerous fields (e.g., business, emergency management, environmental engineering and sciences, geography and spatial sciences, geosciences, public health, and social sciences, to name just a few) and enabled by steady progress of related technologies. As a geospatial data deluge permeates broad scientific and societal realms, to sustain the trend, however, requires innovative cyberGIS capabilities to be developed based on synergistic integration of critical computational and spatial approaches enabled by advanced cyberinfrastructure.

CyberGIS (aka cyber geographic information science and systems based on advanced computing and cyberinfrastructure) has emerged as new-generation GIS comprising a seamless integration of advanced cyberinfrastructure, GIS, and spatial analysis and modeling capabilities while leading to widespread research advances and broad societal impacts. Pioneered by NSF-funded research, cyberGIS has provided a solid foundation for breakthroughs in diverse science, technology and application domains, and contributed to the innovation of cyberinfrastructure overall. During the past several years, cyberGIS has grown as a vibrant interdisciplinary field as evidenced through impactful publications and a number of collaborative projects, meetings, conferences, and workshops. For example, NSF-funded a \$4.8 million multi-institution project: CyberGIS Software Integration for Sustained Geospatial Innovation involves a number of academic institutions, industrial partners (e.g. Esri), U.S. government agency partners (e.g. US Geological Survey), and U.S. federally funded research and development laboratory (e.g., Oak Ridge National Laboratory), and multiple international partners. With an international scope, the project has established a sustainable cyberGIS software framework while achieving major scientific and technological advances in tackling multi-scale environmental and geospatial challenges.

While much progress has been made on cyberGIS innovation and related scientific advances, significant scientific and engineering challenges remain to be solved. For example, in today's geospatially connected world, disasters that occur in any part of the world are rarely isolated events often with impacts cascading through multiple interconnected pathways and systems, which are felt outside the immediately-affected region, in distant parts of the world. Reducing disaster losses and enhancing resilience is a grand challenge nationally and globally as more and more people and assets increasingly are located in hazardous areas. The complexity of these coupled environmental and human systems and their connectivity at various spatial and temporal scales is already a challenging problem in non-disaster times, but becomes critical when time frames for action and response options are constricted during emergencies and disasters. Conventional scientific approaches to understanding such systems and pursuing optimal decisions, however, tend to be fragmented in space and time and constrained by the inability to take advantage of complex, diverse, and massive geospatial data, which make extrapolation over the connectedness across large and multiple spatial and temporal scales difficult or infeasible. Major scientific breakthroughs and synergistic cyberGIS and cyberinfrastructure innovations are urgently needed to discover and understand complex and dynamic geospatial connections between people and places in order to enhance critically important knowledge in this context. An integrated approach combining geospatial big data, analysis and models is expected to ignite transformative geospatial innovation and discovery for enabling effective and timely resolutions of such grand challenges.

Question 2 Cyberinfrastructure Needed to Address the Research Challenge(s) (maximum ~1200 words): Describe any limitations or absence of existing cyberinfrastructure, and/or specific technical advancements in cyberinfrastructure (e.g. advanced computing, data infrastructure, software infrastructure, applications, networking, cybersecurity), that must be addressed to accomplish the identified research challenge(s).

The recent computation- and data-intensive transformation of broad scientific practices has posed significant technical challenges revolving around geospatial data science and cyberGIS. The proliferation of geospatial data sources (e.g., remote sensing, sensor networks, unmanned aerial vehicles, and social media) have brought up new challenges such as poor or unknown data quality, missing metadata, and lack of well-defined sampling schemes. The proliferation of these data streams represents a pressing big data problem that is not likely to slow down in the foreseeable future, thus innovation of advanced cyberinfrastructure-enabled geospatial data and software capabilities is required to address it. This innovation opportunity is significant and unique because geospatial analysis and modeling workflows often simultaneously involve computing-, data-, and visualization-intensive capabilities with responsive user interactivity expected and, thus, cannot be effectively supported by conventional cyberinfrastructure. Major scientific and technological breakthroughs need to be pursued through holistic approaches to geospatial integration with advanced cyberinfrastructure and cyberGIS.

Based on such approaches, the ROGER supercomputer (<http://go.illinois.edu/roger/>) funded by the NSF MRI program has made significant progress on establishing a generic advanced cyberinfrastructure resource with collaborative, data-intensive, high-performance, and

Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-05 16:38:07

PAGE 3

REFERENCE NO: 292

interactive capabilities employed to solve significant problems in cross-cutting, bio, engineering, geo, and social sciences that would otherwise be impossible or difficult to solve. For the first time as a national cyberinfrastructure resource made available as part of the XSEDE program, ROGER has achieved the integration of HPC/GPU, data-intensive computing, and cloud computing as a holistic hardware and software environment. While demonstrating the feasibility and importance of this integration through ROGER, more important challenges and opportunities have been identified with a set highlighted as follows, which are deemed critical to future innovation of advanced cyberinfrastructure and cyberGIS.

* Interactive computing providing users' instant access to cyberinfrastructure power will empower geospatial discovery and innovation. Oversubscribed resources with long queue times reduces the usability of cyberinfrastructure for a range of use cases. Spatial decision support systems (SDSSs), for example, are used in defense, intelligence, federal and local governments, agriculture, industry, and science alike. Yet, decision makers and domain scientists are reaching a point where they cannot cope with the growing amounts of data using traditional SDSSs without research breakthroughs. To address these needs requires a new breed of undersubscribed cyberinfrastructure resources supporting interactive computing with close coupling between web servers and compute nodes. These new systems are expected to better support interactive decision making and collaborative problem solving.

* Innovating cyberinfrastructure and cyberGIS for pursuing science and engineering frontiers includes many dimensions. Users often access cyberinfrastructure through web services and web-based user interfaces. cyberinfrastructure resources must support multimodal access, especially online access, through flexible policies and computational resource configurations, which should be straightforward to develop and deploy in safe and secure ways. Further, to meet the needs of an advancing field with rapid software develop cycles requires streamlined policies that support customized software environments. Flexible policies that allow developers rather than system admins to resolve package dependencies and software installations will likely support this rapid development cycle.

* Data movement to, from, and within cyberinfrastructure is an impediment for many users new to this paradigm, which prohibits broader, widespread use and adoption of cyberinfrastructure. Increased network bandwidth and user-friendly technologies such as Globus.org help to alleviate this challenge, but streamlined policies, technologies, and approaches such as cyberGIS may help users bring data to cyberinfrastructure to leverage advanced computing and visualization technologies.

* Standardized cyberGIS workflows for large geospatial data repositories are important to operational programs. For example cyberGIS workflows for elevation, drainage, and other feature extractions from lidar, high resolution satellite images, and other geospatial sources including social media are needed.

* Research and development of semantics for geospatial big data is a significant domain challenge. For example, extracting and providing rich semantic representations for geospatial features from big data such as lidar and high resolution images represent a range of computation- and data-intensive problems.

The synergistic cyberinfrastructure and cyberGIS advances will positively reverberate into broad scientific advances as evidenced by a number of initiatives and projects sponsored by various agencies and organizations (e.g., NSF, EPA, NASA, USDA, and USGS) in the evolution of cyberGIS and geospatial data science. Furthermore, major impacts on geospatial technologies and geo-services are desirable to be pursued through partnership with industry – an industry estimated in 2013 to have between \$150 billion and \$270 billion annual global revenue. Transformative advances in cyberGIS and geospatial data science education efforts will likely produce a much-needed workforce for fostering cyberinfrastructure-enabled discoveries and innovations. New fundamental understanding promises to be gained for many important societal problems (e.g., critical infrastructures, emergency management, national security).

Consent Statement

• "I hereby agree to give the National Science Foundation (NSF) the right to use this information for the purposes stated above and to display it on a publically available website, consistent with the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>)."